Simulations / statistical mechanics / …

Andrew Torda, April 2010

Topics

- very simple stat mechanics
- Monte Carlo
- molecular dynamics

Favourite books

- "Understanding Molecular Simulation", Frenkel and Smit, Academic Press, 2002
- "Computer Simulation of Liquids", Allen & Tildesley, Oxford Science Publications, 1990

Misc background

- simpler than last year
    - 2^{nd} law mostly removed

Plan for today

- 45 min lecture

- 45 Übung / lecture
    - how to calculate a force

Concepts

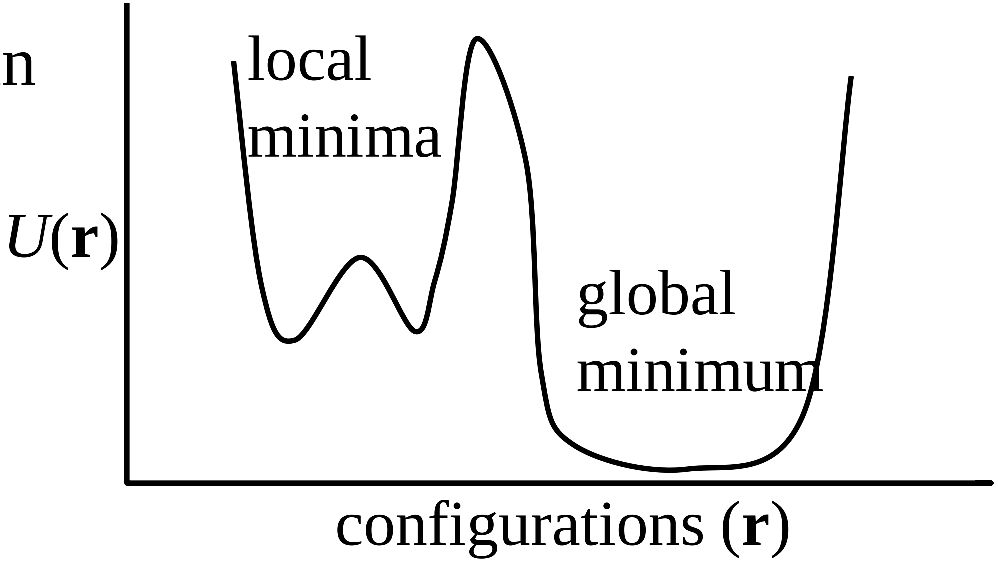
- potential energy
- work
- entropy
- free energy

Nomenclature

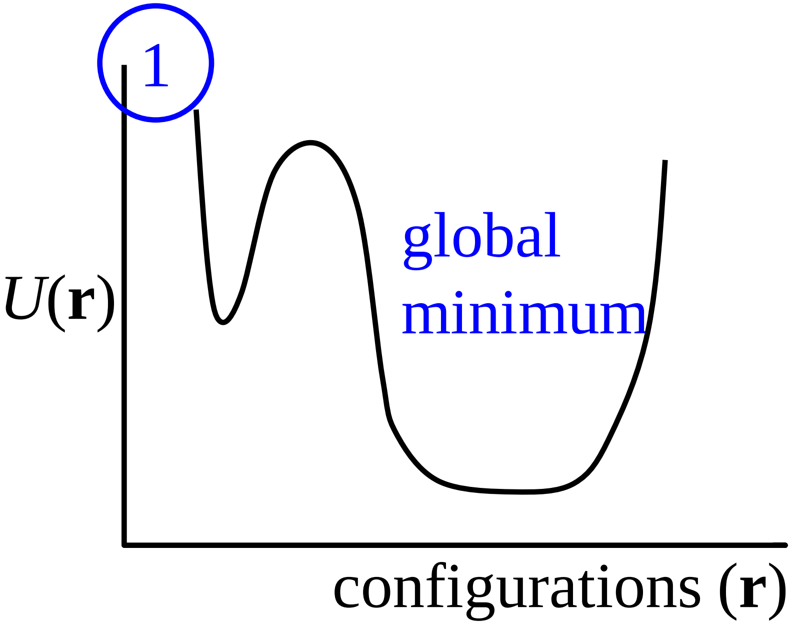
- T temperature
- N all kinds of things, usually number of particles
- k and k_B Boltzmann's constant
- R gas constant = $k\ N_a$
- S entropy
- F Helmholtz (NVT)
- G (Gibbs) free energy (NPT)
- U internal energy
- H enthalpy = $U + pV$
- E energy, E_{kin}, E_{pot}, E_{bond}, …
- Q heat
- W work
- V volume
- Ω (x) number of states of system with property (x)

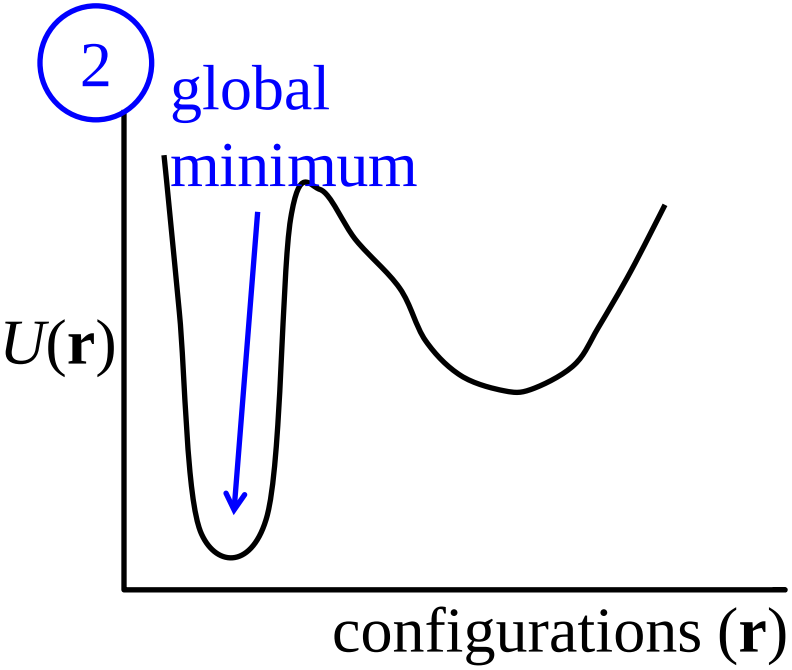
Potential energy

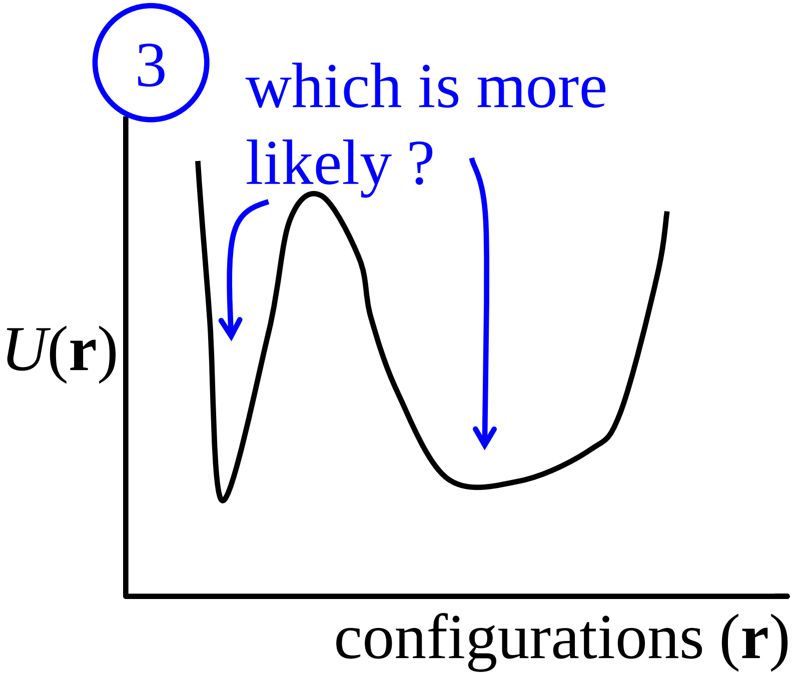
- electrostatic $U(r) = \frac{q_1 q_2}{4\pi\varepsilon_0 r}$
- gravity $U(r) = \frac{G m_1 m_2}{r}$
- elastic, …
- potential energy of a protein…
  - where would the protein like to be ?



More than potential energy

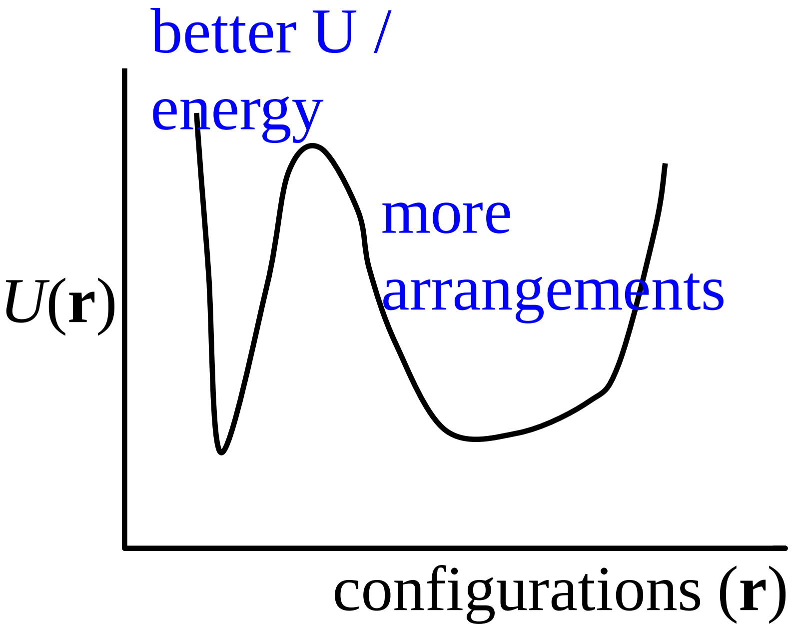






- in more detail …

Simulation and formal statistical mechanics



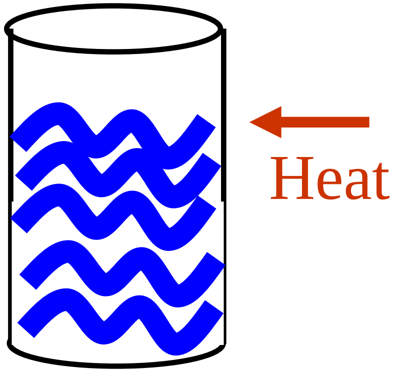
Quantified ?

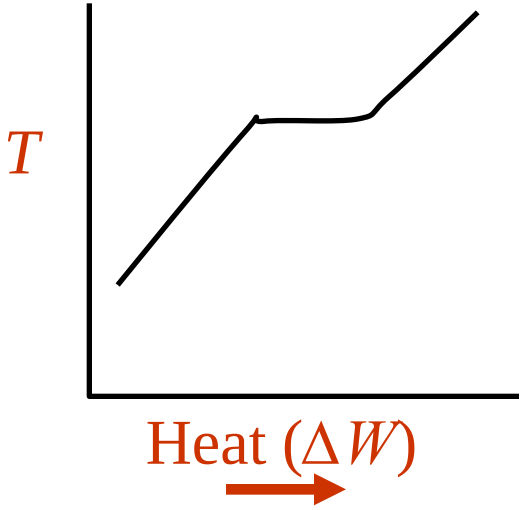
- entropy and free energy

Dumb simulation, preferring to go downhill

- should show how probabilities (entropy) are balanced against energy

Heat Capacity example





- change in rate of heating ? Boiling / phase change
- proteins ? DNA ? folding / melting

- easy to simulate ?
  - only if certain rules are followed
    - (example) no heat can leave our system
- Important
  - simulations are valid if they follow rules

History

- Statistical mechanics
  - derived by summing up properties of individual particles
- thermodynamics
  - less emphasis on individual particles
- lots of formulae which cannot be completely applied to proteins
  - sums over infinite volumes, numbers of particles, time

Rules and limitations

- Always at equilibrium

First Law

Conservation of energy

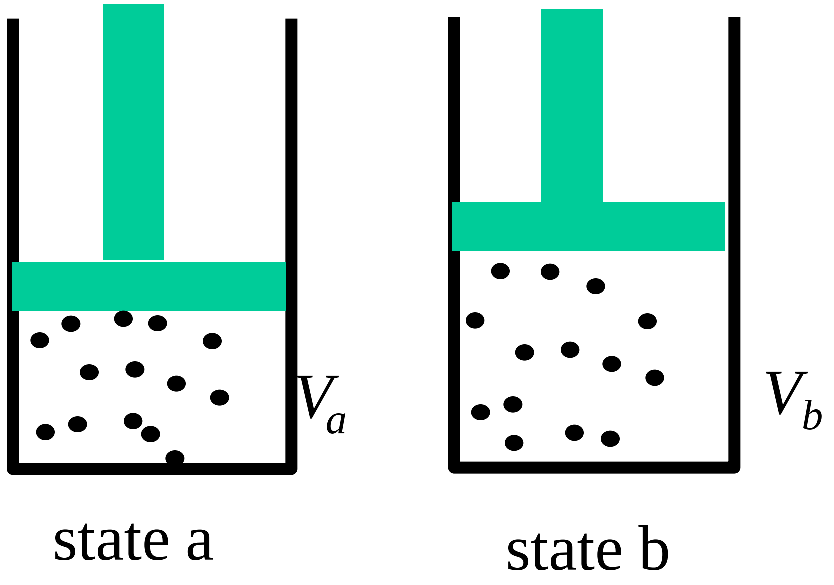
- $dU = \Delta Q + \Delta W$
- dU change in internal energy
- ΔQ heat given to system
- ΔW work done on system

- example of work…

Work on a gas

- $\Delta W = -P\,\Delta V$

- $dU = \Delta Q + \Delta\, W$
  $= \Delta Q - P\,\Delta V$



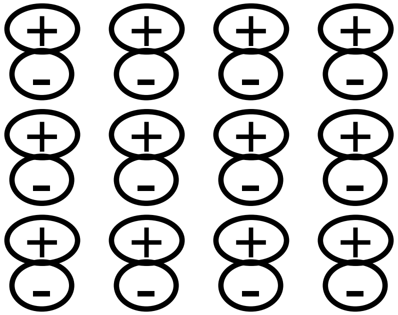
- others
  - charges in electric fields, surface tension / area, ….

Entropy

- second law of thermodynamics
  - you tend to become disordered
- role in free energy
- formula for simple systems

- Disorder, how many ways can the system arrange itself..
  - depends on energy (and volume and number of particles)
- How many states can the system have / occupy (specified E) ?
  - $\Omega(E)$
  - $S = k \ln \Omega(E)$

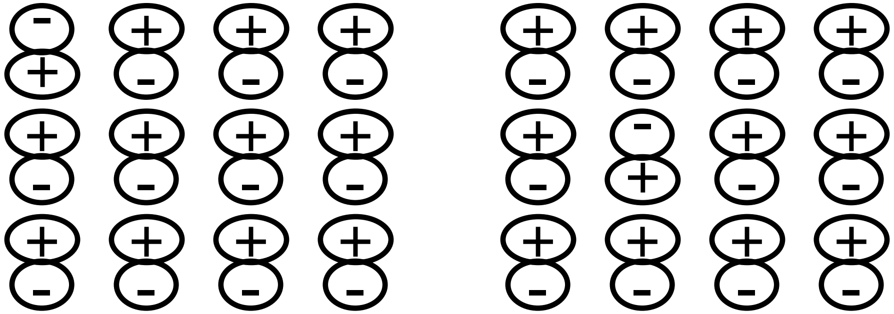
Number of states $\Omega(X)$



T=0

$\Omega(E) = 1$

$S = -k \ln 1 = 0$



….. $T > 0$ small

$\Omega(E) = N_{atoms}$

$S = k \ln N_{atoms}$

- a bit more energy ? more states
  - more.. solid →liquid .. many many more

Gibbs

What if states are not equally likely ?

- N_{state} states with distribution
- p_1=0.999, p_2=0.00001, p_3=0.00001, ….
- just as if the system had one state
  - low entropy
- p_1=0.5, p_2=0.1, p_3=0.1, ….
  - a bit more entropy, but still very much dominated by p_1
- p_1=0.01, p_2=0.01, p_3=0.01, ….
  - lots of states, all equally likely
  - lots of entropy

In general

$$S = -k \sum_{i=1}^{N_{state}} p_i \ln p_i$$

Gibbs entropy !!

Applicability

- can one really estimate ? $S = -k \sum_{i=1}^{N_{state}} p_i \ln p_i$
  - liquid.. no
  - protein ..no
- with a simulation ?
  - too many states
  - approximations to ΔS later
- simple system (grid / lattice)
  - yes .. later

Second law

- disorder increases

$$dS = \frac{dQ_{rev}}{T} \geq \frac{dQ}{T}$$

S entropy
Q heat

- in an irreversible process, S of system always increases
- useful consequence, for small changes
- $dQ = TdS$

- Intuitive ?
  - I heat the system, temperature does not go up much
  - making water boil

not so important

Entropy and other properties

- if $dQ = TdS$

$$\begin{aligned} dU &= dQ + dW \\ &= TdS + dW \\ &= TdS - PdV \end{aligned}$$

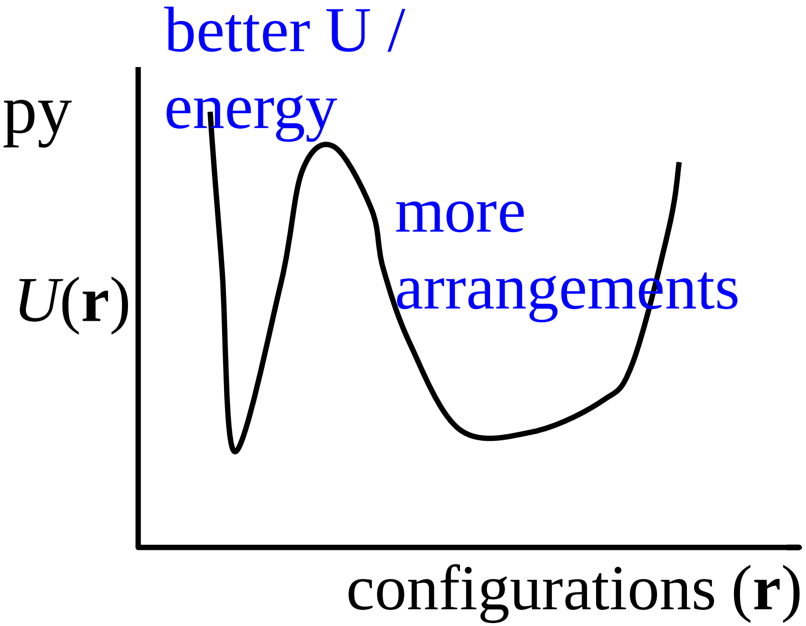
- what if we fix volume ? $\left(\dfrac{dU}{dS}\right)_V = T$

- Typical technique..
  - fix V or P or T and look at the relations
  - fixed (N, V, T) Helmholtz
  - fixed (N, V, P) Gibbs

not so important

Units

- more arrangements = more entropy
- behaving like energy
- entropy units
  - energy / temp (JK^{-1})



- from picture, we do not care about *U*, introduce *F* (Helmholtz)
- $F = U - TS$

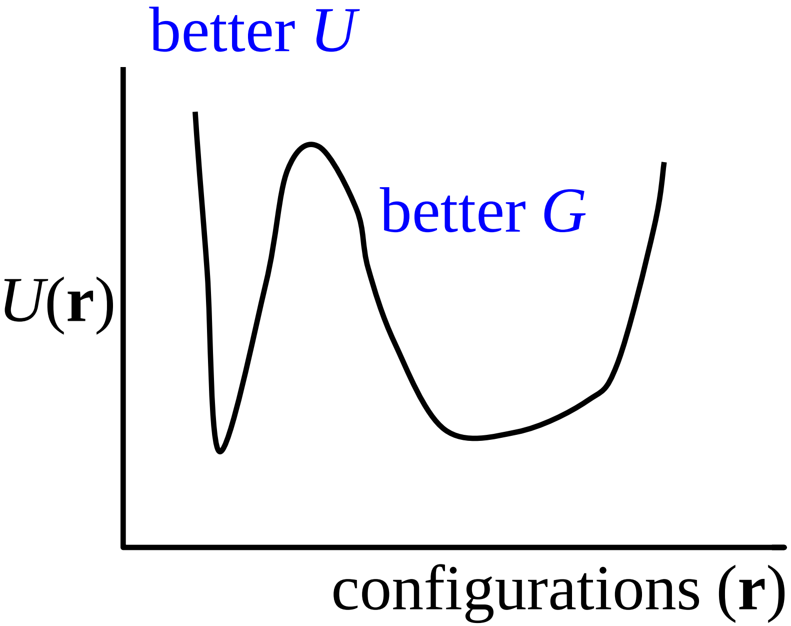
Free Energy types

- $F = U - TS$
- $G = U - TS + pV$

  $= H - TS$

- often we look at changes
- $\Delta G = H - T\,\Delta S$

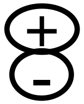
- is my picture valid ?
  - not strictly (what if system can hop around ?)



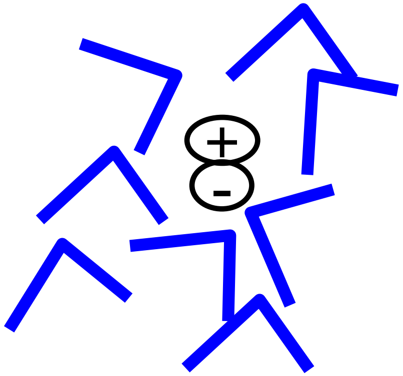
not so important

Where can we use this ?

- what is the entropy of a molecule ?
  - sounds easy

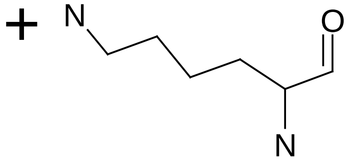


  - really…
  - entropy depends on solvent (system)



- entropy of a conformation ?
  - has no meaning

$$S = -k \sum_{i=1}^{N_{state}} p_i \ln p_i$$



- is it additive ?
  - sometimes

Adding entropy

Ω_{A} Ω_{B} $\Omega_{\mathrm{AB}} = \Omega_{\mathrm{A}}\,\Omega_{\mathrm{B}}$

- entropy is extensive
  - $S_{AB} = -\,k \ln(\Omega_{\mathrm{A}}\,\Omega_{\mathrm{B}}) = -\,(k \ln(\Omega_{\mathrm{A}}) + k \ln(\Omega_{\mathrm{B}})) = S_A + S_B$
- assumption
  - for my new system A and B weakly interact
- what if they interact ?
  - putting A in state 1 changes probability of B in state 1
  - what if it just changes the probability ?

$\Omega_{\mathrm{AB}} \neq \Omega_{\mathrm{A}}\,\Omega_{\mathrm{B}}$

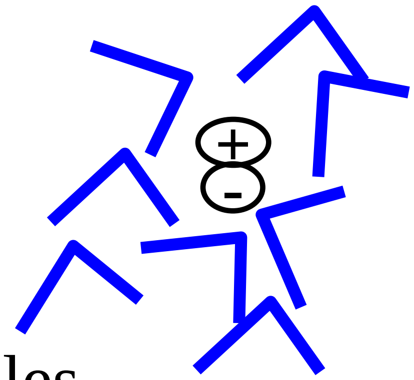
$$S = -k \sum_{i=1}^{N_{state}} p_i \ln p_i$$

Adding entropy

- I cannot usually add the entropy of two systems to get the entropy of new total system

- can I decompose entropy ?

May we decompose entropy ?



- Remember energy in proteins
- $E_{tot} = E_{bonds} + E_{vdw} + E\ldots$
- $S_{tot} = S_{bonds} + S_{vdw} + \ldots$?
  - no
  - makes no sense unless
  - bonds are decoupled from atoms and angles…
- Different parts of system
  - $S_{ligand} + S_{solvent} + \ldots$
- obviously they do interact
- Free energies
  - G_{solv}, $G_{protein}$, G_{ligand}, G_{bonds} not really legal

Decomposing Free Energy

- Would you see this in the literature ?

- Example
  - in protein X asp → asn mutation (COO− to CON)
  - protein becomes more stable

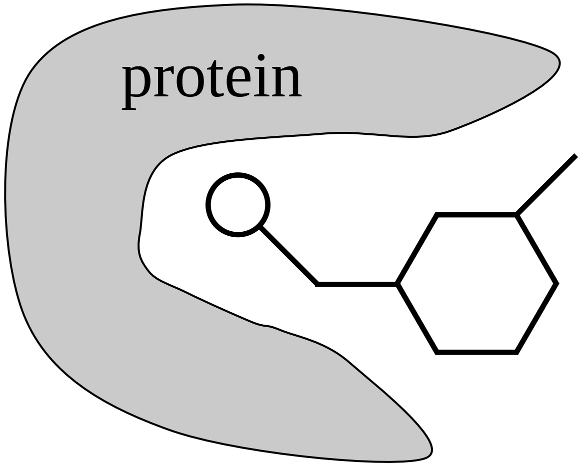
  - stability depends on $\Delta G_{\mathrm{fold} \rightarrow \mathrm{misfold}}$
  - the protein is more stable due to a change in electrostatic free energy

More examples – particle interactions

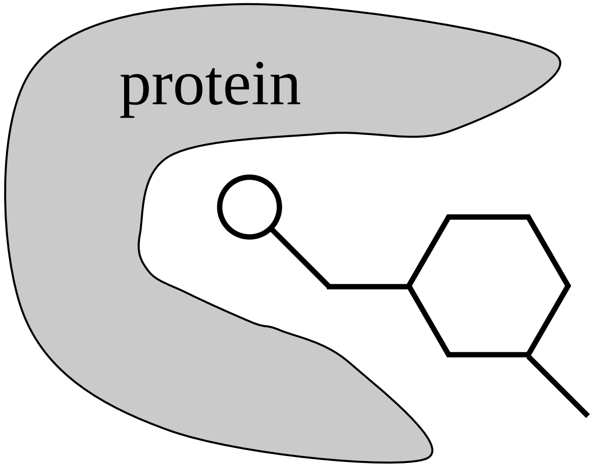
- You have a scoring function for interactions
- usually called ΔG
    - $\Delta G_{tot} = \Delta G_{HB} +$ H-bonds
      $\Delta G_{ionic} + \ldots.$ ionic

- why is this bad nomenclature ?

Entropy of one conformation ?



binding 1



binding 2

- can I talk about ΔG_1 vs ΔG_2 ?
  - $S = - k \ln \Omega$
    - but Ω depends on all accessible states
- if binding 1 can change to binding 2, they are both part of Ω
- I do believe we can talk about U_1 and U_2

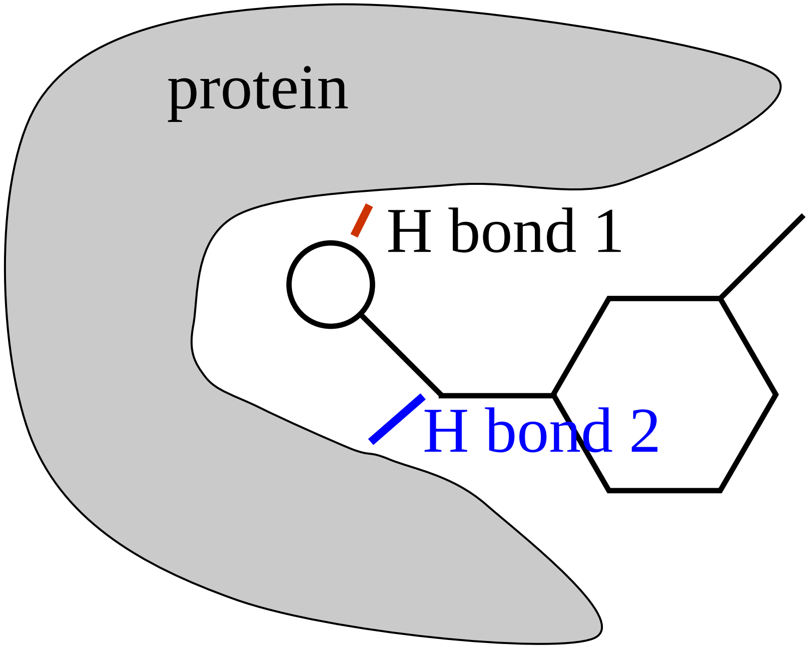
Meaning of labelled free energies

- if we have "free energy contributions"
  - we have entropy contributions
- ΔG_{HB} implies $H_{HB} + \Delta\ TS_{HB}$
- ΔG_{ionic} implies $H_{ionic} + \Delta\ TS_{ionic}$
  - what is S_{ionic}? $\mathrm{S} = -\ k \ln \Omega_{\mathrm{ionic}}$
    - no meaning

- one cannot have "free energy contributions"

independence of terms

- even without labelled entropies



- can we add free energy due to Hbond 1 and 2 ?
- implies $\Delta S = \Delta S_1 + \Delta S_2$
- $-k \ln \Omega = -(k \ln \Omega_1 + k \ln \Omega_2)$
- $-k \ln \Omega = -k \ln (\Omega_1 \Omega_2)$

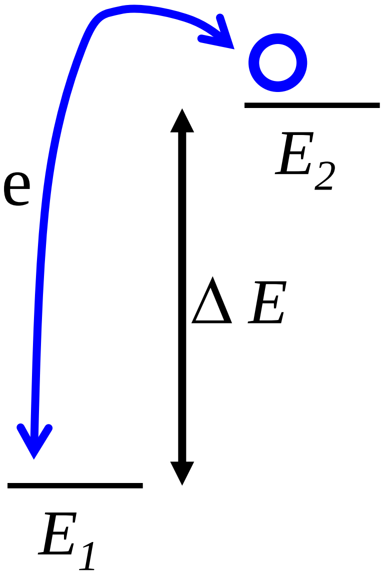
- only possible if there is no interaction between 1 and 2

- additivity of free energies is widely used (usually wrong)

Stop and summarise

- Internal energy U – things like springs
- Enthalpy H – includes pressue
- Gibbs free energy G what we usually use
- Free energy lets us incorporate the concept of what is most likely
- Entropy is not additive between systems that interact

Boltzmann distribution – two states



- System with two energy levels
- I know the temperature and energy difference
- How likely is system to sit in E_1 ? E_2 ?
  - System is very cold,
    - ΔE seems big
  - System is very very hot T→∞
    - ΔE does not matter, seems small
- Relative populations (probabilities) p_1, p_2
- $\Delta E = E_2 - E_1$

$$\frac{p_1}{p_2} = e^{\Delta E / kT}$$

- sometimes

$$\frac{p_1}{p_2} = e^{\beta \Delta E}$$

Boltzmann distribution

- what is the probability of a certain energy level ?
- depends on all available levels

$$p_i = \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}}$$

- name of bottom … partition function, Z

$$Z = \sum_i e^{-E_i/kT}$$

- does this agree with previous slide ?

Consequences of Boltzmann distribution

- At absolute zero
  - only lowest energy state is populated
- At low temperatures
  - low energy states favoured
- High temperature
  - system can visit high energy regions
- Infinite temperature
  - all states equally likely
- For two states (bound / unbound)
  - exponential term means populations quickly become big/small

$$\frac{p_1}{p_2} = e^{\Delta E / kT}$$

How big are energy differences ?

$$\frac{p_1}{p_2} = e^{\Delta E / kT}$$

- simplest case, equal populations
- $p_1 = p_2$

$$1 = e^{\Delta E / kT}$$

$$\ln 1 = \Delta E / kT$$

$$\Delta E = 0$$

- k=1.38×10^{-23} J K^{-1}, but use
- $R = k\, N_A$

  $= k$ 6.02×10^{23} J K^{-1}mol^{-1}

  = 8.314 J K^{-1} mol^{-1}

examples of populations

- for 99:1 at 300 K
  - ΔE=11 kJ mol^{-1}

$$\frac{p_1}{p_2} = e^{\Delta E / kT}$$

$$\ln \frac{p_1}{p_2} = \Delta E / kT$$

$$\Delta E = kT \ln \frac{p_1}{p_2}$$

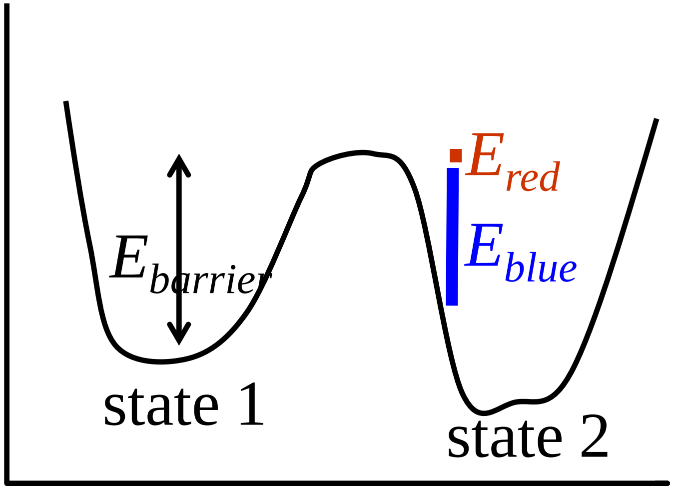
Drugs

- Dissociation constant k_d of drug D to protein P
- nanomolar drugs are 10^{-9},
- difference between 1 and 10 nm binding drug
  - 46 vs 52 kJ mol^{-1}
- topic will return later

$$k_d = \frac{[\mathrm{D}][\mathrm{P}]}{[\mathrm{DP}]}$$

Barrier crossing

- How likely are you to cross a barrier ?
  - p_{red} vs p_{blue}



- all the blue copies of system will not make it over $E_{barrier}$
- red population is small, even for $E_{barrier} \sim kT$
- but explains why $\ln(rate) \propto T$

where next ?

- how to see some of these properties by simulating
- how simulating depends on these properties